AFATL-TR-71-113

AD735274

GENERALIZATIONS OF A CLASS OF FREQUENCY FUNCTIONS FOR SYSTEMS ANALYSIS

AIR TO AIR MISSILES AND TARGETS DIVISION

TECHNICAL REPORT AFATL-TR-71-113

SEPTEMBER 1971

Approved for public release; distribution unlimited.

AIR FORCE ARMAMENT LABORATORY

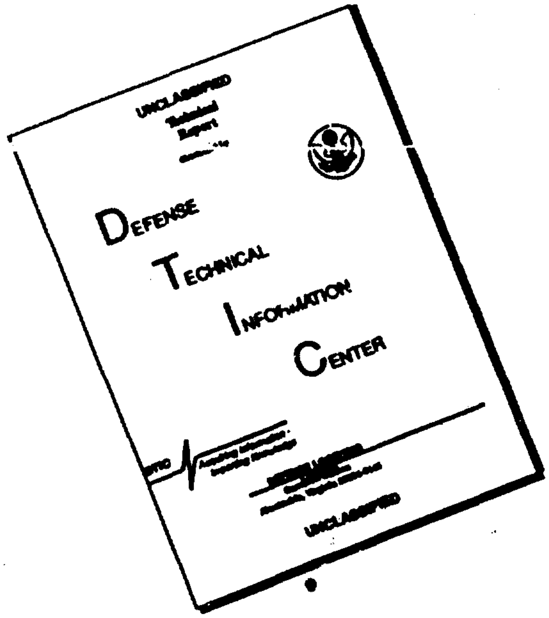
AIR FORCE SYSTEMS COMMAND • UNITED STATES AIR FORCE

EGLIN AIR FORCE BASE, FLORIDA

K

94

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Air to Air Missiles and Targets Division Air Force Armament Laboratory Eglin Air Force Base, Florida 32542	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

GENERALIZATIONS OF A CLASS OF FREQUENCY FUNCTIONS FOR SYSTEMS ANALYSIS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Final Report - September 1971

5. AUTHOR(S) (First name, middle initial, last name)

Robert N. Braswell
T. Clark Pewitt

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1971	44	4

8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) AFATL-TR-71-113 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Approved for Public Release; Distribution Unlimited.

11. SUPPLEMENTARY NOTES Available in DDC	12. SPONSORING MILITARY ACTIVITY Air Force Armament Laboratory Air Force Systems Command Eglin Air Force Base, Florida 32542

13. ABSTRACT

A family of generalized models are presented for analyzing data or systems with stochastic properties. Previous work in this field is presented as well as a new finite range frequency distribution function with nomograms, tables, examples, and recent extensions to a generalized family of methods and models. The analytical techniques are related to statistical mechanics and were developed specifically for analyses of weapon systems.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Frequency function Systems analysis Statistical mechanics Distribution function						

Generalizations of a Class of Frequency Functions for Systems Analysis

Robert N. Braswell

T. Clark Pewitt

FOREWORD

This report is not based on a specific Air Force project but represents information that is of potential value to the scientific community.

This technical report has been reviewed and is approved.

William W. Metz

WILLIAM W. METZ, Chief
Air to Air Missiles and Targets Division

ABSTRACT

A family of generalized models are presented for analyzing data or systems with stochastic properties. Previous work in this field is presented as well as a new finite range frequency distribution function with nomograms, tables, examples, and recent extensions to a generalized family of methods and models. The analytical techniques are related to statistical mechanics and were developed specifically for analyses of weapon systems.

TABLE OF CONTENTS

(The reverse of this page is blank)

SECTION I

INTRODUCTION

In solving problems involving stochastic processes, the systems analyst and engineer often have to choose a probability distribution function that will fit the experimental data. One method is to use one of the typical "text-book" distributions. These are generally tabulated, with established computational procedures for finding moments. This approach has proven useful, especially when higher moments of the distribution are needed. However in many cases, (a) only the first moment, or the first and second moments, are needed in the analysis, (b) there may be obvious similarities to given data and curves encountered in previous situations, and (c) there may be an advantage in using a standard distribution that will allow ease in some predetermined calculations and tests. In such cases, it may be desirable to devise distribution functions that fit the analysis, rather than encumbering the analysis with unnecessarily complicated functions and fitting procedures.

In this report, a class of distribution functions is derived and it is shown that the frequency function developed by Braswell and extended by Manders, References 1 and 2 (see the Appendix), is a particular member of a family of these distribution functions. Procedures and methods of finding parameters of these distributions are presented. To show the simplicity and usefulness of this method, the parameters of the distribution of Braswell and Manders are re-derived, using the more general techniques. Finally an example of application of these types of distributions is given.

SECTION II

CLASS OF FREQUENCY FUNCTIONS

Suppose there is given a function $h(x)$, such that

$$\frac{dh(x)}{dx} > 0 \qquad \text{for } x_0 \leq x \leq x_1 \tag{1}$$

Then the doubly truncated cumulative distribution function

$$F(x) = \begin{cases} 0 \text{ for } x \leq x_0 \\ \dfrac{h(x) - h(x_0)}{h(x_1) - h(x_0)} & \text{for } x_0 < x < x_1 \\ 1 \text{ for } x_1 \leq x \end{cases} \tag{2}$$

is a proper cummulative distribution function (cdf) with finite domain associated with positive probability.

The corresponding probability distribution function is given by

$$f(x) = \frac{dF(x)}{dx} = \begin{cases} 0 \text{ for } x \notin [x_0, x_1] \\ \dfrac{h'(x)}{h(x_1) - h(x_0)} & \text{for } x \varepsilon [x_0, x_1] \end{cases} \tag{3}$$

The mean, μ, is given by

$$\mu = \int_{x_0}^{x_1} xf(x)\, dx \tag{4}$$

$$= xF(x) \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} F(x)\, dx \tag{5}$$

Since $F(x_0) = 0$ and $F(x_1) = 1$,

$$\mu = x_1 - \int_{x_0}^{x_1} F(x)\, dx. \tag{6}$$

Similarly, the second moment, μ_2, is given by,

$$\mu_2 = x^2F(x) - 2\int_{x_0}^{x_1} xF(x)\,dx \qquad 7$$

Since

$\frac{d}{dx}\int_{x_0}^{x} F(t)\,dt = F(x)$, Equation 7 may be written,

$$\mu_2 = x^2F(x)\Big|_{x_0}^{x_1} - 2x\int_{x_0}^{x} F(t)\,dt\Big|_{x_0}^{x_1} + 2\int_{x_0}^{x_1}\int_{x_0}^{x} F(t)\,dtdx \qquad 8$$

Therefore, a useful form for μ_2 if F(x) is twice integrable is

$$\mu_2 = x_1^2 - 2x_1\int_{x_0}^{x_1} F(t)dt + 2\int_{x_0}^{x_1}\int_{x_0}^{x} F(t)\,dtdx \qquad 9$$

Using Equations 6 and 9, μ_2 may be expressed by

$$\mu_2 = x_1\mu - x_1\int_{x_0}^{x_1} F(t)dt + 2\int_{x_0}^{x_1}\int_{x_0}^{x} F(t)dtdx \qquad 10$$

The absolute deviation, γ, is given by

$$\gamma = \int_{x_0}^{x_1} |x - u|f(x)\,dx \qquad 11$$

For $x_0 \geq 0$,

$$\gamma = -\int_{x_0}^{\mu} xf(x)\,dx + \mu\int_{x_0}^{\mu} f(x)\,dx + \int_{\mu}^{x_1} xf(x)\,dx - \mu\int_{\mu}^{x_1} f(x)\,dx \qquad 12$$

$$= -xF(x)\Big|_{x_0}^{\mu} + \int_{x_0}^{\mu} F(x)\,dx + \mu[F(\mu) - F(x_0)] + xF(x)\Big|_{\mu}^{x_1}$$

$$-\int_{\mu}^{x_1} F(x)\,dx - \mu[F(x_1)-F(\mu)] \qquad 13$$

$$= \int_{x_0}^{\mu} F(x)\,dx - \int_{\mu}^{x_1} F(x)\,dx - \mu + x_1 \qquad 14$$

To show the simplicity of use, the form and parameters of the frequency function of Braswell and Manders are derived.

Here, $h(x) = \dfrac{de^{ax}}{1 + be^{ax}}$ for $x_0 = 0$ and $x_1 = 1$ 15

and, thus $h'(x) > 0$, for $a > 0$, $x \varepsilon [0, 1]$.

Also,

$$h(x_0) = h(0) = \frac{d}{1 + b}$$

$$h(x_1) = h(1) = \frac{de^a}{1 + be^a}$$

Therefore,

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \dfrac{\left(\dfrac{e^{ax}}{1 + be^{ax}}\right) - \left(\dfrac{1}{1 + b}\right)}{\left(\dfrac{e^{a}}{1 + be^{a}}\right) - \left(\dfrac{1}{1 + b}\right)} = H(x) & \text{for } 0 \leq x \leq 1 \\ 1 & \text{for } x > 1.0 \end{cases} \qquad 16$$

To get F(x) into the form previously used, we examine H(x).

$$H(x) = \frac{(e^{ax} - 1)(1 + be^{a})}{(1 + be^{ax})(e^{a} - 1)}$$

$$= \frac{[1 + e^{-\alpha(1+2\delta)}]}{[1 - e^{-2\alpha}]} \quad \frac{[1 - e^{-2\alpha x}]}{[1 + e^{\alpha(1-2x-2\delta)}]} \quad \text{for } 0 \leq x \leq 1 \qquad 17$$

where $\alpha = -\dfrac{a}{2}$ and $\delta = \dfrac{a + 2 \ln b}{2a}$

Multiplying Equation 17 by $\dfrac{e^{\alpha(1-2\delta)}}{e^{\alpha(1-2\delta)}}$

$$H(x) = \frac{[1 + e^{-\alpha(1+2\delta)}][e^{\alpha(1-2\delta)} - e^{\alpha(1-2x-2\delta)}]}{[e^{\alpha(1-2\delta)} - e^{-\alpha(1+2\delta)}][1+e^{\alpha(1-2x-2\delta)}]} \quad \text{for } 0 \leq x \leq 1 \qquad 18$$

which is the result obtained by Braswell and Manders.

The mean, μ, is given by Equation 6, therefore

$$\mu = xF(x) \Big|_0^1 - \int_0^1 F(x)\ dx \qquad 19$$

Thus, μ may be written as

$$\mu = 1 - \left(\frac{1 + be^a}{1 - e^a}\right) \int_0^1 \frac{1 - e^{ax}}{1 + be^{ax}}\ dx \qquad 20$$

The integral in Equation 20 is evaluated in Reference 1.

Due to desired generalizations for the cdf of Braswell and Manders presented later in this paper, it is desirable to be able to evaluate integrals of the form

$$I = \int_{x_0}^{x_1} \frac{e^{cx} dx}{b+e^{ax}} \qquad 21$$

Thus, the transformation, $\eta = e^{ax}$ gives

$$I = \frac{1}{a} \int_{e^{ex_0}}^{e^{ax_1}} \frac{\eta^{c/a-1} d\eta}{b+\eta} \qquad 22$$

For $b > - e^{ax_0}$,

$$\frac{1}{\eta^{1-c/a}(b + \eta)} = \frac{1}{b} \frac{1}{\eta^{1-c/a}} - \frac{1}{b} \frac{\eta^{c/a}}{(b + \eta)} \qquad 23$$

Therefore,

$$I = \frac{1}{ab} \int_{e^{ax_0}}^{e^{ax_1}} \frac{dx}{x^{1-c/a}} - \frac{1}{ab} \int_{e^{ax_0}}^{e^{ax_1}} \frac{x^{c/a} dx}{(b + x)} \qquad 24$$

Evaluating the first integral,

$$I_1 = \frac{1}{ab} \int_{e^{ax_0}}^{e^{ax_1}} \frac{dx}{x^{1-c/a}} = \begin{cases} \dfrac{1}{bc}\, x^{c/a} \Big|_{e^{ax_0}}^{e^{ax_1}} & \text{for } c \neq 0 \\ \dfrac{1}{ab}\, \ln x \Big|_{e^{ax_0}}^{e^{ax_1}} & \text{for } c = 0 \end{cases} \qquad 25$$

Therefore;

$$I_1 = \begin{cases} \frac{1}{bc} \; [e^{ax_1} - e^{ax_0}] & \text{for } c \neq 0 \\ \\ \frac{1}{b} \; [x_1 - x_0] & \text{for } c = 0 \end{cases} \qquad 26$$

The second integral to be considered is,

$$I_2 = \frac{1}{ab} \int_{e^{ax_0}}^{e^{ax_1}} \frac{x^{c/a}}{(b+x)} \, dx \qquad 27$$

If $c = 0$, then

$$I_2 = \frac{1}{ab} \int_{e^{ax_0}}^{e^{ax_1}} \frac{dx}{b+x} \qquad 28$$

or

$$I_2 = \frac{1}{ab} \, \ln \left[\frac{e^{ax_1}}{e^{ax_0}} \right] \qquad 29$$

If $c/a = n$, some positive integer, then I_1 is given by Equation 25, and

$$I_2 = \frac{1}{ab} \int_{e^{ax_0}}^{e^{ax_1}} \frac{x^n}{b+x} \; dx \qquad 30$$

$$= \frac{1}{ab} \int_{e^{ax_0}+b}^{e^{ax_1}+b} \frac{(x-b)^n}{x} \, dx$$

$$= \frac{1}{ab} \int_{e^{ax_0}+b}^{e^{ax_1}+b} \sum_{j=0}^{n} \binom{n}{j} x^j \; (-b)^{n-j} \frac{dx}{x} \qquad 31$$

$$I_2 = \frac{(-b)^n}{ab} \int_{e^{ax_0}+b}^{e^{ax_1}+b} \frac{dx}{x} + \frac{(-b)^n}{ab} \sum_{j=1}^{n} \binom{n}{j} \; (-b)^{-j} \int_{e^{ax_0}+b}^{e^{ax_1}+b} x^{j-1} dx \qquad 32$$

$$I_2 = \frac{(-b)^n}{ab} \ln \left[\frac{e^{ax_1}+b}{e^{ax_0}+b} \right] + \frac{(-b)^n}{ab} \sum_{j=1}^{n} \binom{n}{j} (-b)^{-j} \left[\frac{\left(e^{ax_1}+b\right)^j - \left(e^{ax_0}+b\right)^j}{j} \right] \qquad 33$$

If c/a is not a positive integer or 0, then I_2 may be written in the form

$$I_2 = \frac{(-b)^{c/a}}{ab} \sum_{j=0}^{\infty} \binom{c/a}{j} (-b)^{-j} \int_{e^{ax_0}+b}^{e^{ax_1}+b} x^{j-1}\, dx \qquad 34$$

where $\binom{x}{k}$ is defined for non-integral x as, k a non-negative integer, as

$$\binom{x}{k} = \begin{cases} 0 & \text{for } k = 0 \\ x(x-1)\ldots(x-k+1) & \text{for } k \geq 1 \end{cases} \qquad 35$$

Therefore

$$I_2 = \frac{(-b)^{c/a}}{ab} \ln \left[\frac{e^{ax_1}+b}{e^{ax_1}+b} \right] + \frac{(-b)^{c/a}}{ab} \sum_{j=1}^{\infty} \binom{c/a}{j} \left(\frac{1}{b}\right)^j \frac{\left(e^{ax_1}+b\right)^j - \left(e^{ax_0}+b\right)^j}{j} \qquad 36$$

It should be noted in Equation 36 that if $b > 0$ then the expression $(-b)^{c/a}$ has both real and imaginary components, but the imaginary part may be neglected in this case. The method may be used to evaluate integrals such as.

(a) $\displaystyle\int_{x_0}^{x_1} \frac{\sinh(\alpha x)}{\sinh(\beta x)}\, dx$

(b) $\displaystyle\int_{x_0}^{x_1} \frac{\cosh(\alpha x)}{\sinh(\beta x)}\, dx$

(c) $\displaystyle\int_{x_0}^{x_1} \frac{\cosh(\alpha x)}{\cosh(\beta x)}\, dx$

and many other similar forms. The evaluation of these integrals is simplified when the ratio α/β is a positive integer (the smaller the integer,

the fewer number of terms necessary in the evaluation).

An immediate result of the foregoing discussion is a method of evaluating slightly more general integrals, those of the form

$$I = \int_{x_0}^{x_1} \frac{e^{cx}dx}{\left(b+e^{ax}\right)^{\alpha}} \tag{37}$$

Here

$$I = \frac{1}{a}\int_{e^{x_0}}^{e^{x_1}} \frac{x^{c/a}}{x(b+x)^{\alpha}}\, dx \tag{38}$$

$$= \frac{1}{a}\int_{e^{x_0}+b}^{e^{x_1}+b} \frac{(x-b)^{c/a-1}}{x^{\alpha}}\, dx \tag{39}$$

$$= \frac{1}{a}\sum_{j=0}^{\infty} \binom{c/a-1}{j} (-b)^{c/a-1-j} \int_{e^{x_0}+b}^{e^{x_1}+b} x^{j-\alpha}\, dx \tag{40}$$

Again, if c/a is a positive integer, there will be only $c/a + 1$ terms in Equation 40.

It is of interest to note that the foregoing integrals are generalizations of many of the tabulated integrals. Therefore,the evaluation of these integrals is of some importance in itself.

Returning to the determination of the mean and absolute deviation of the distribution of Braswell and Manders, the mean is obtained using Equations 20, 26, and 31. The absolute deviation is obtained by use of Equations 14, 26, and 31. When $\alpha = -\frac{a}{2}$ and $\delta = \frac{a+2\ \ell nb}{2a}$, the results in Reference 1 are obtained.

For completeness, the frequency function of Braswell and Manders should be examined for the limiting cases:

(a) $\lim\limits_{a\to 0_+} F(x)$ (b) $\lim\limits_{b\to -1_+} F(x)$

For case (a)

$$\lim_{a\to 0_+} F(x) = \lim_{a\to 0_+} \left(\frac{e^{ax}-1}{e^{a}-1}\right)\left(\frac{1+be^{a}}{1+be^{ax}}\right) \quad 41$$

If $b \neq -1$, then

$$\lim_{a\to 0_+} F(x) = \lim_{a\to 0_+} \left(\frac{e^{ax}-1}{e^{a}-1}\right) \quad 42$$

Therefore, by L'Hospital's rule

$$\lim_{a\to 0_+} F(x) = x \quad 43$$

Thus F(x) tends to the cdf of a random variable uniformly distributed in the interval {0,1}.

In case (b)

$$\lim_{b\to -1_+} F(x) = \left(\frac{e^{ax}-1}{e^{a}-1}\right) \lim_{b\to -1_+} \left(\frac{1+be^{a}}{1+be^{ax}}\right) \quad 44$$

or

$$\lim_{b\to -1_+} F(x) = 1 \quad 45$$

Thus x is a variable associated with a deterministic process.

A useful extension of the frequency function of Braswell and Manders given in Reference 1 is given by

$$h(x) = \frac{e^{cx}}{b + e^{ax}} \quad 46$$

with $x_o = 0$, $x_1 = 1$, $b > -1$.

For monotonicity of F(x), $\frac{dh(x)}{dx}$ must be greater than zero, $\forall x$ for $x \in [0,1]$.

Since $\frac{dh(x)}{dx} = \frac{ce^{cx}}{b+e^{ax}} - \frac{ae^{cx}e^{ax}}{(b+e^{ax})^2}$, the condition 47

$\frac{dh(x)}{dx} > 0$ will hold if $c > \frac{ae^{ax}}{b+e^{ax}}$ $\forall x$ for $x \in [0,1]$

Since $w(x) = \dfrac{e^{ax}}{b+e^{ax}}$ is monotone increasing for $x \in [0,1]$, the maximum of $w(x)$ occurs at $x = 1$, and therefore, monotonicity is insured if $c > \dfrac{ae^{a}}{b+e^{a}}$, ($b \geq 0$ and $c \geq a$ will insure monotonicity).

It should be noted that h(x) given in Equation 46 is a generalization of the energy distribution functions of statistical mechanics as given in Reference 3. The Maxwell-Boltzmann distribution is of the form,

$$f(x) = \frac{1}{e^{x} e^{x/kt}} \qquad 48$$

The Bose-Einstein distribution is of the form

$$f(x) = \frac{1}{e^{a} e^{x/kt} - 1} \qquad 49$$

and finally, the Fermi-Dirac distribution is given by

$$f(x) = \frac{1}{e^{a} e^{x/kt} + 1} \; . \qquad 50$$

Using h(x) given in Equation 49, the cdf takes the form

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \dfrac{\left(\dfrac{e^{cx}}{b+e^{ax}}\right) - \left(\dfrac{1}{b+1}\right)}{\left(\dfrac{e^{c}}{b+e^{a}}\right) - \left(\dfrac{1}{b+1}\right)} & \text{for } 0 \leq x \leq 1 \\ 1 & \text{for } x > 1 \end{cases} \qquad 51$$

The function given in Equation 15 is a particular example of a family of functions given by

$$h(x) = \frac{de^{cx}}{(b+e^{ax})^{2}} \qquad \text{for } x_0 = 0 \text{ and } x_1 = 1 \qquad 52$$

Using this function, the corresponding cdf is given by

$$F(x) = \begin{cases} 0 \quad \text{for } x \leq 0 \\ \dfrac{\left(\dfrac{e^{cx}}{(b+e^{ax})^{\alpha}}\right) - \left(\dfrac{1}{b+1}\right)^{\alpha}}{\left(\dfrac{e^{c}}{(b+e^{a})^{\alpha}}\right) - \left(\dfrac{1}{b+1}\right)^{\alpha}} \quad \text{for } 0 \leq x \leq 1 \\ 1 \quad \text{for } x > 0 \end{cases} \tag{53}$$

To evaluate the parameters of this distribution, integrals of the form

$$I = \int_{e^{\beta}}^{e^{\gamma}} \frac{e^{cx}}{(b+e^{ax})^{\alpha}} \, dx \tag{54}$$

must be evaluated and the evaluation of these integrals is greatly simplified (thus increasing the usefulness of the distribution) if $c/a = n$, a positive integer and $\alpha = k$, some positive integer.

In this case, Equation 54 becomes

$$I = 1/a \int_{e^{\beta}}^{e^{\gamma}} \frac{x^{n-1}}{(b+x)^{k}} \, dx \tag{55}$$

$$= 1/a \int_{e^{\beta+b}}^{e^{\gamma+b}} \frac{1}{x^{k}} \sum_{j=0}^{n-1} \binom{n-1}{j} x^{j} (-b)^{n-1-j} dx \tag{56}$$

$$= \frac{(-b)^{n-1}}{a} \sum_{j=0}^{n-1} \binom{n-1}{j} (-b)^{-j} \int_{e^{\beta+b}}^{e^{\gamma+b}} x^{j-k} \, dx \tag{57}$$

There are four parameters associated with this distribution: a, b, c, and k (n is determined by a and c) and formulae have been developed earlier for evaluating the mean and absolute deviation. Therefore, if any two of the parameters are fixed, the mean and absolute deviation can be used to determine the other two. By doing this, six families of distribution functions are generated.

For example, if $b = 1$, $a = 1$, $x_0 = 0$, $x_1 = 1$, then

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \dfrac{\left(\dfrac{e^{nx}}{(1+e^{x})^{k}}\right) - \left(\dfrac{1}{2k}\right)}{\left(\dfrac{e^{n}}{(1+e)^{k}}\right) - \left(\dfrac{1}{2k}\right)} & \text{for } 0 \leq x \leq 1 \\ 1 & \text{for } x > 0 \end{cases} \qquad 58$$

SECTION III

RANDOM NUMBERS

Often, it is necessary to use random numbers selected from a distribution. The cdf of Braswell and Manders may be easily inverted to give random numbers from that distribution if random variables uniformly distributed in {0,1} are available.

If F(x) is monotone increasing cdf, and r is a variable from a uniform {0,1} distribution, then

$$y = F^{-1}(x)$$

is a random number from F(x).

For the FRPDF of Braswell and Manders, the cdf may be inverted in closed form, thus making generation of corresponding random numbers easy.

Let

$$F(x) = r = \frac{\left(\frac{e^{ax}}{1+be^{ax}}\right) - \left(\frac{1}{1+b}\right)}{\left(\frac{e^{a}}{1+be^{a}}\right) - \left(\frac{1}{1+b}\right)}$$

Then

$$F^{-1}(x) = x = \frac{1}{a}\, \ell n \frac{r\left(\frac{e^{a}}{1+be^{a}} - \frac{1}{1+b}\right) + \left(\frac{1}{1+b}\right)}{1 - b\left[r\left(\frac{e^{a}}{1+be^{a}} - \frac{1}{1+b}\right) + \left(\frac{1}{1+b}\right)\right]}$$

Random numbers using the function by Braswell and Manders are given in Table I.

For the generalizations of the above function, the inverse cannot be found in closed form. This presents no real problems, in that "Monte-Carlo" techniques in Reference 4 may be used to generate accurately as many random numbers from these distributions as may be needed in any problem that is to be analyzed.

TABLE I. RANDOM NUMBERS USING THE FRPDF OF BRASWELL AND MANDERS

0.24158	0.60196	0.25900	0.97604	0.81010
0.32882	0.54244	0.02165	0.99494	0.35506
0.57602	0.21315	0.83320	0.37908	0.46429
0.46475	0.47708	0.07142	0.19321	0.46652
0.59312	0.59904	0.55056	0.37447	0.75369
0.74804	0.00798	0.53492	0.38361	0.84148
0.19745	0.06575	0.11968	0.55770	0.64208
0.42256	0.74125	0.65361	0.77012	0.59652
0.93307	0.38003	0.21965	0.22837	0.02158
0.32717	0.51644	0.43901	0.31461	0.38016
0.73331	0.88885	0.23166	0.54247	0.09798
0.58225	0.18766	0.31569	0.79382	0.04368
0.26919	0.25216	0.43510	0.26635	0.00619
0.50120	0.67743	0.35515	0.32886	0.48071
0.96026	0.04537	0.41694	0.79842	0.68748
0.37894	0.50911	0.43165	0.87127	0.67873
0.99545	0.08539	0.21373	0.58254	0.36961
0.74520	0.19261	0.47631	0.79148	0.19704
0.19635	0.29674	0.40224	0.36312	0.51801
0.75558	0.82781	0.99734	0.18026	0.00213
(A=0.10000, B=0.90000)				

Example:

It is illustrative to show how the generalized class of frequency functions can be used in systems analysis.

A sample of times to complete jobs on a computer was taken (Table II). A useable empirical distribution was needed to fit such samples.

To handle this problem, the probability distribution associated with the generalized cdf was analyzed.

It was noted that these jobs fell into two categories. The first category was short jobs, i.e., jobs that took less than one minute to complete. The other category was long jobs, i.e., those that used more than one minute to complete.

The shorter jobs far outnumbered the longer jobs at the installation from which the data was taken. Also, it was known that the longer jobs were usually "production runs" of tested programs, so that they could be handled differently than the short test jobs.

Thus the data of interest concerned the shorter running programs. The associated pdf, f(x) may be written,

$$f(x) = \frac{\left[\dfrac{e^{ax}}{(b+e^{ax})^k}\right]\left[c - \dfrac{kae^{ax}}{b+e^{ax}}\right]}{\left[\dfrac{e^{c}}{(b+e^{a})^k} - \dfrac{1}{(b+1)^k}\right]} \qquad 59$$

Since $\dfrac{e^{ax}}{(b+e^{ax})^k}$ is an increasing function of x, and $c - \dfrac{kae^{ax}}{b+e^{ax}}$ is a decreasing function of x, it was seen that the important middle range values of the sample could be fitted by varying some of the parameters.

Since, only means and absolute deviations were easily computed, the cdf given by Equation 58 was adopted. Then

$$f(x) = \frac{\left[\dfrac{e^{nx}}{(1+e^{x})^k}\right]\left[n - \dfrac{ke^{x}}{1+e^{x}}\right]}{\left[\dfrac{e^{n}}{(1+e)^k} - \dfrac{1}{2^k}\right]}$$

A table of means and absolute deviations for various n and k is shown in Table II and used to pick a value of n and k to agree with the estimated mean and absolute deviation of the sample, which were $\hat{\mu} = .2534$, and $\hat{\gamma} = .17$.

TABLE II. μ AND γ AS A FUNCTION OF n AND k

K	n = 1 μ	n = 1 γ	k	n = 2 μ	n = 2 γ
0	0.577	0.235	0	0.652	0.219
1	0.475	0.242	1	0.598	0.235
2	0.640	0.201	2	0.506	0.240
3	0.494	0.238	3		
4	0.429	0.227	4	0.624	0.200
5	0.378	0.217	5	0.497	0.229
6	0.333	0.206	6	0.428	0.225
7	0.296	0.188	7	0.374	0.217
8	0.263	0.174	8	0.329	0.196
9	0.236	0.156	9	0.291	0.189
	n = 3			**n = 4**	
0	0.174	0.189	0	0.764	0.161
1	0.669	0.204	1	0.729	0.175
2	0.613	0.228	2	0.688	0.197
3	0.536	0.237	3	0.635	0.218
4	0.383	0.212	4	0.566	0.233
5			5	0.457	0.228
6	0.608	0.197	6		
7	0.494	0.228	7		
8	0.423	0.224	8	0.592	0.204
9	0.369	0.207	9	0.487	0.220
	n = 5			**n = 6**	
0	0.802	0.138	0	0.831	0.119
1	0.776	0.151	1	0.811	0.132
2	0.744	0.171	2	0.787	0.146
3	0.705	0.190	3	0.758	0.159
4	0.657	0.208	4	0.722	0.183
5	0.594	0.228	5	0.678	0.198
6	0.504	0.233	6	0.620	0.223
			7	0.548	0.232
			8	0.420	0.207

For the given sample, $n = 1$ and $k = 8$ were chosen. To test the goodness-of-fit of the fitted distribution, the data were divided in five group and an χ^2 test performed. See raw data in Tables III and IV.

The regions considered were (data in Table V).

I: $0 \leq x < 0.1$

II: $0.1 \leq x < 0.2$

III: $0.2 \leq x < 0.3$

IV: $0.3 \leq x < 0.4$

V: $0.4 \leq x \leq 1.0$

Setting $\chi^2 = \sum_{m=1}^{5} \frac{(y_m - g_m)^2}{g_m}$, where y_m = number of data points in interval m, and g_m = expected number of points in interval m. g_m was found using the tabulated value of F(x), and χ^2 was found to be 4.087. Since there were five intervals, and two estimated parameters, the variable χ^2 should belong to a chi-squared distribution with two degrees of freedom. The value of 4.087 indicates a significance level of about 85 percent.

Much higher values of χ^2 are found if a normal distribution is hypothesized, due to the fact that $\sigma^2 = .044$. This small value of σ^2 causes g_1 and g_5 to be small, thus increasing the value of χ^2. The given χ^2 for this problem could be greatly reduced if the restriction of integral n and k were removed, or if larger values of integral n and k were considered.

CONCLUSION

The scheme presented in this paper may be extended to use for different "basis" functions for distributions. To use this method the analyst should:

(a) Determine if there is something to be gained by using this technique. If the problem under consideration is not made easier, he should not pursue this technique.

TABLE III. SEQUENCE OF COLLECTED DATA

Observation	Time	Observation	Time	Observation	Time	Observation	Time
1	0.46	12	0.05	23	0.14	34	0.04
2	0.23	13	0.22	24	0.02	35	0.08
3	0.35	14	0.39	25	0.36	36	0.02
4	0.10	15	0.13	26	0.26	37	0.53
5	0.37	16	0.88	27	0.02	38	0.04
6	0.10	17	0.31	28	0.08	39	0.40
7	0.07	18	0.08	29	0.10	40	0.13
8	0.34	19	1.00	30	0.05	41	0.23
9	0.09	20	0.49	31	0.47	42	0.26
10	0.59	21	0.15	32	0.39	43	0.30
11	0.32	22	0.27	33	0.12	44	0.02

TABLE IV. ORDERED DATA

Order	Time	Order	Time	Order	Time	Order	Time
1	0.02	12	0.02	23	0.02	34	0.02
2	0.04	13	0.04	24	0.05	35	0.05
3	0.07	14	0.08	25	0.08	36	0.08
4	0.09	15	0.10	26	0.10	37	0.10
5	0.12	16	0.13	27	0.13	38	0.14
6	0.15	17	0.22	28	0.23	39	0.23
7	0.26	18	0.26	29	0.27	40	0.30
8	0.31	19	0.32	30	0.34	41	0.35
9	0.36	20	0.37	31	0.39	42	0.39
10	0.40	21	0.46	32	0.47	43	0.49
11	0.53	22	0.59	33	0.88	44	1.00

TABLE V. CLUSTERED DATA

Interval	No. of Observations
0 to 0.1	13
0.1 to 0.2	8
0.2 to 0.3	6
0.3 to 0.4	9
0.4 to 1.0	8

Parameter estimates: $\hat{\mu} = 0.25$, $\hat{\gamma} = 0.17$, $\hat{\sigma} = 0.044$

(b) Pick a family of functions that may be parameterized, and that he can manipulate in some way.

(c) Determine the forms of the associated cdf's and pdf's. Of particular interest would be critical points of these functions.

(d) Establish a method (tables of means, deviations, etc.; explicit formulas; etc.) of estimating the parameters of these distributions.

Once he has these tools he may then:

(e) Proceed to solve the problem at hand.

Objections to this type of use may be raised, and the analyst would have to determine the appropriateness of applying this method in a particular application. One point to be considered in this determination, is that, in general problems that are presented for solution are so large or complicated that a digital computer must be used. If a numerical approach is used for parts of a problem, then the analyst should be able to use this method if it makes some parts of the problem more manageable. But, since this is the only case in which this method should be considered, most objections to the technique can be overcome.

(The reverse of this page is blank)

APPENDIX

Much of the information in this report is based on mathematical equations presented in References 1 and 2. Those papers were published only in Japan and are not readily available in the United States. Therefore, by special arrangement with the publisher, the papers are included here as an appendix.

(The reverse of this page is blank)

A NEW FINITE RANGE PROBABILITY DISTRIBUTION FUNCTION (FRPDF) WITH PARAMETERS-NOMOGRAM AND TABLES*

By Robert N. BRASWELL
University of Florida
Gainesville, Florida

Clara Fu-Mei MANDERS
Radiation, Incorporated Melbourne, Florida

The objective of this paper is to introduce a flexible finite range probability distribution function, FRPDF. The emphasis is on flexibility of application relative to its simplicity of use and its ability to fit varied experimental data clusters.

The FRPDF, $f(x)$, $0 \leq x \leq 1$, is strictly unimodal as shown in Figures 2 through 7. When $\delta = 0$, the probability density function is always symmetrical with respect to the vertical line $x = 0.5$, and the distribution function is a family of S-shaped curves. When $\delta > 0$, the peak of the probability density function shifts toward the $x = 1$ line and when $\delta > 0$, it shifts to the $x = 0$ line. When $\alpha = 0$, the probability density function is identically uniform. For the large value of α, $\alpha = 10$ or more the peak of the probability density function becomes narrower and higher at appropriate values of x for given values of δ. Then, δ can be considered as the location parameter, while α can be considered as the shaping parameter and these parameters will give the desirable feature; namely, flexibility to the probability density function.

1. Introduction

The purpose of developing this new distribution function is to fulfill the need for a flexible and easy-to-use finite range probability distribution function for many experimental problems. Knowledge of the probability distribution of a random variable is required before statistical inferences can be made. All real random variables, by their very nature, have a finite range. The function used as a starting point in this paper was originally used by Pearl to approximate the population growth characteristic of the United States[1]. Then Braswell[2] reformulated it into Task Operating Characteristics (TOC) curves. This paper transforms the TOC curve into hyperbolic expressions as shown in Equation (4) and into the finite probability density function shown in Equation (5).

The developed FRPDF is like the beta distribution function in that it contains two parameters, α and δ, and that the admissible values of the variate X lie between zero and one. Unlike the beta distribution, it is flexible and simple to apply as illustrated in this paper. In general, this distribution is useful in any application requiring statistical analysis of collected experimental data. The computer was used to develop tables, curves, and a

* Received 24, Feb., 1970.
This work was supported in Part by THEMIS Contract DAH C04 68C 0002.
(Reprinted from Rep. Stat. Appl. Res., JUSE, Vol 17, No. 2, 1970, by permission of the publisher.)

Preceding page blank

nomograph to relate the mean and the absolute deviation of the sample with the density function parameters, α and δ.

2. Mathematical Basis

Pearl's function for the population growth of the United States is defined as:

$$y = \frac{be^{ax}}{1+ce^{ax}}, \qquad x \geq 0 \text{ (time)} \tag{1}$$

where y = number of people; a, b, and c are parameters; and e is the base for Naperian logarithms. Braswell's first paper changed Pearl's function into the following form:

$$P(t) = \frac{[1+e^{-\alpha(1+2\delta)}][e^{\alpha(1-2\delta)}-e^{\alpha(1-2t-2\delta)}]}{[e^{\alpha(1-2\delta)}-e^{-\alpha(1+2\delta)}][1+e^{\alpha(1-2t-2\delta)}]} \tag{2}$$

for $0 \leq t \leq 1$, $\alpha > 0$, and $-0.5 \leq \delta \leq +0.5$.

Since $P(t)$ is not defined when $\alpha = 0$, it is necessary to define this function at this point by use of L 'Hospitals' rule.

$$\lim_{\alpha \to 0} P(t) = \lim_{\alpha \to 0} \frac{\frac{d}{d\alpha} N(\alpha, \delta, t)}{\frac{d}{d\alpha} D(\alpha, \delta, t)} = t$$

where

$$N(\alpha, \delta, t) = [1+e^{-\alpha(1+2\delta)}][e^{\alpha(1-2\delta)}-e^{\alpha(1-2t-2\delta)}]$$

and

$$D(\alpha, \delta, t) = [e^{\alpha(1-2\delta)}-e^{-\alpha(1+2\delta)}][1+e^{\alpha(1-2t-2\delta)}].$$

For more completeness, Equation (2) becomes

$$P(t)\begin{cases} = \dfrac{[1+e^{-\alpha(1+2\delta)}][e^{\alpha(1-2\delta)}-e^{\alpha(1-2t-2\delta)}]}{[e^{\alpha(1-2\delta)}-e^{-\alpha(1+2\delta)}][1+e^{\alpha(1-2t-2\delta)}]} & \text{for } \alpha>0 \\ = t \quad \text{for } \alpha=0 & 0\leq t\leq 1. \end{cases} \tag{3}$$

3. Formulation of the FRPDF

For convenience change the notation of the independent variable from t to x and P to F, and let $F(x)$ be a probability distribution function. When one rewrites these equations more explicitly, one obtains the following result. For $\alpha>0$, $P(t)$ becomes:

$$F(x) = \frac{[1+e^{-\alpha(1+2\delta)}][e^{\alpha(1-2\delta)}-e^{\alpha(1-2x-2\delta)}]}{[e^{\alpha(1-2\delta)}-e^{-\alpha(1+2\delta)}][1+e^{\alpha(1-2x-2\delta)}]}$$

and for $\alpha=0$ $F(x)=x$, for $0\leq x\leq 1.0$.

Equation (3) can be rewritten as follows:

$$F(x)\begin{cases} = \dfrac{\cosh[\alpha(\delta+0.5)]}{\sinh\alpha}\,\dfrac{\sinh\alpha x}{\cosh[\alpha(x+\delta-0.5)]} & \text{for } \alpha>0 \\ = x \quad \text{for } \alpha=0 & 0\leq x\leq 1. \end{cases} \tag{4}$$

Since $F(x)$ has the following properties, it is a valid probability distribution function[3]:

(a) $F(0)=0$, $F(1)=1$

(b) It is a nondecreasing function of x:

$$F(x_1) \leq F(x_2) \qquad \text{for } x_1 < x_2$$

(c) It is continuous from the right:

$$F(x^+) = F(x)$$

i.e., if $x_1 < x_2$, then $P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1)$.

For a continuous random variable, the probability density function is $F'(x)$. For $\alpha>0$,

$$f(x)=\frac{dF(x)}{dx}$$

$$=\left(\frac{\alpha\cosh[\alpha(\delta+0.5)]\cosh[\alpha(\delta-0.5)]}{\sinh\alpha}\right)\left(\frac{1}{\cosh^2[\alpha(x+\delta-0.5)]}\right).$$

Consequently the complete form for the FRPDF is:

$$f(x)\begin{cases}=G(\alpha,\ \delta)\dfrac{1}{\cosh^2[\alpha(x+\delta-0.5)]} & \text{for } \alpha>0 \quad 0\leq x\leq 1\\ =1.0 \quad \text{for } \alpha=0 \quad 0\leq x\leq 1.0 & \\ =0 \text{ elsewhere} & \end{cases} \tag{5}$$

where

$$G(\alpha,\ \delta)=\frac{\alpha\cosh[\alpha(\delta+0.5)]\cosh[\alpha(\delta-0.5)]}{\sinh\alpha}.$$

The useful parameter approximations of the FRPDF are the mean and absolute deviation of the samples. The mean and the absolute deviation as a function of α and δ is derived. It will also be shown that the variance is finite and that the moment generating function exists.

4. The Mean

The mean, μ, is by definition:

$$\mu=E(x)=\int_{-\infty}^{\infty}xf(x)dx. \tag{6}$$

Substitute the Equation (5) into Equation (6), for $\alpha>0$, $0\leq x\leq 1$,

$$\mu(\alpha,\ \delta)=\int_0^1 xG(\alpha,\ \delta)\frac{1}{\cosh^2[\alpha(x+\delta-0.5)]}dx$$

$$=G(\alpha,\ \delta)\int_0^1\frac{x}{\cosh^2[\alpha(x+\delta-0.5)]}dx.$$

Let $z=\alpha(x+\delta-0.5)$.

Then μ becomes

$$\mu(\alpha,\ \delta)=G(\alpha,\ \delta)\left[\int_{\alpha(\delta-0.5)}^{\alpha(\delta+0.5)}\frac{z}{\alpha\cosh^2 z}\frac{dz}{\alpha}+\int_{\alpha(\delta-0.5)}^{\alpha(\delta+0.5)}(0.5-\delta)\frac{1}{\cosh^2 z}\frac{dz}{\alpha}\right]$$

$$=G(\alpha,\ \delta)\left[\frac{1}{\alpha^2}\int_{\alpha(\delta-0.5)}^{\alpha(\delta+0.5)}\frac{z}{\cosh^2 z}dz+\frac{(0.5-\delta)}{\alpha}\int_{\alpha(\delta-0.5)}^{\alpha(\delta+0.5)}\frac{1}{\cosh^2 z}dz\right].$$

From integral tables [4],

$$\mu(\alpha,\ \delta)=G(\alpha,\ \delta)\frac{1}{\alpha^2}\left[\alpha\tanh[\alpha(\delta+0.5)]+\log\frac{\cosh[\alpha(\delta-0.5)]}{\cosh[\alpha(\delta+0.5)]}\right]$$

for $\alpha>0$ and

$$\mu=\int_0^1 xdx=0.5 \qquad \text{for } \alpha=0.$$

The mean is therefore

$$\mu(\alpha,\ \delta)=\begin{cases}=\dfrac{G(\alpha,\ \delta)}{\alpha^2}\left[\alpha\tanh[\alpha(\delta+0.5)]+\log\dfrac{\cosh[\alpha(\delta-0.5)]}{\cosh[\alpha(\delta+0.5)]}\right] & \text{for } \alpha>0\\ =0.5 \quad \text{for } \alpha=0 & \end{cases} \tag{7}$$

and $G(\alpha,\ \delta)$ is defined in Equation (5).

5. Absolute Deviation [5]

The absolute deviation, denoted γ, is defined as

$$\gamma(\alpha,\ \delta)=\int_{-\infty}^{\infty}|x-\mu|\,f(x)dx \tag{8}$$

Table 1 Mean as a function of α and δ

α \ δ	-0.5	-0.4	-0.3	-0.2	-0.1	0.0	+0.1	+0.2	+0.3	+0.4	+0.5
1	0.5696	0.5570	0.5436	0.5295	0.5149	0.5000	0.4851	0.4705	0.4564	0.4430	0.4304
2	0.6872	0.6533	0.6257	0.5870	0.5445	0.5000	0.4555	0.4130	0.3743	0.3407	0.3128
3	0.7736	0.7385	[illegible]	[illegible]	0.5693	0.5000	0.4307	0.3655	0.3052	0.2615	0.2264
4	0.8274	0.7888	0.7337	0.6641	0.5845	0.5000	0.4155	0.3359	0.2663	0.2112	0.1726
5	0.8615	0.8205	0.7538	0.6807	0.5926	0.5000	0.4074	0.3135	0.2412	0.1795	0.1385
6	0.8845	0.8414	0.7740	[illegible]	0.5965	0.5000	0.4035	0.3105	0.2260	0.1586	0.1155
7	0.9010	0.8557	0.7834	[illegible]	0.5984	0.5000	0.4016	0.3055	0.2166	0.1443	0.0990
8	0.9134	0.8660	0.7872	[illegible]	0.5993	0.5000	0.4007	0.3036	0.2107	0.1340	0.0866
9	0.9230	0.8736	0.7930	0.6934	0.5997	0.5000	0.4003	0.3016	0.2070	0.1264	0.0770
10	0.9307	0.8743	0.7954	0.6951	0.5999	0.5000	0.4001	0.3009	0.2046	0.1207	0.0693
11	0.9370	0.8836	0.7970	0.6995	0.5999	0.5000	0.4001	0.3005	0.2030	0.1164	0.0630
12	0.9422	0.8870	0.7930	0.6987	0.6000	0.5000	0.4000	0.3003	0.2020	0.1130	0.0578
13	0.9467	0.8896	0.7937	[illegible]	0.6000	0.5000	0.4000	0.3001	0.2013	0.1104	0.0533
14	0.9505	0.8917	0.7931	[illegible]	0.6000	0.5000	0.4000	0.3001	0.2009	0.1083	0.0495
15	0.9538	0.8933	0.7994	0.7000	[illegible]	0.5000	0.4000	0.3000	0.2006	0.1067	0.0462
16	0.9567	0.8946	0.7996	0.7000	[illegible]	0.5000	0.4000	0.3000	0.2004	0.1054	0.0433
17	0.9592	0.8957	0.7997	0.7000	0.6000	0.5000	0.4000	0.3000	0.2003	0.1043	0.0408
18	0.9615	0.8965	0.7998	0.7000	0.6000	0.5000	0.4000	0.3000	0.2002	0.1035	0.0385
19	0.9635	0.8972	0.7993	[illegible]	0.6000	0.5000	0.4000	0.3000	0.2001	0.1028	0.0365
20	0.9653	0.8977	0.7939	0.70	.6000	0.5000	0.4000	0.3000	0.2001	0.1025	0.0347

Table 2 Absolute deviation as a function of α and δ

α \ δ	-0.5	-0.4	-0.3	-0.2	-0.1	0.0	+0.1	+0.2	+0.3	+0.4	+0.5
1	0.2361	0.2374	0.2385	0.2394	0.2399	0.2401	0.2399	0.2394	0.2385	0.2374	0.2361
2	0.1910	0.1978	0.2044	0.2101	0.2139	0.2152	0.2139	0.2101	0.2044	0.1978	0.1910
3	0.1439	0.1545	0.1658	0.1758	0.1826	0.1850	0.1826	0.1758	0.1658	0.1545	0.1439
4	0.1109	0.1229	0.1359	0.1469	0.1540	0.1564	0.1540	0.1469	0.1359	0.1229	0.1109
5	0.0892	0.1017	0.1148	0.1249	0.1306	0.1324	0.1306	0.1249	0.1148	0.1017	0.0892
6	0.0744	0.0871	0.0998	0.1081	0.1121	0.1132	0.1121	0.1081	0.0998	0.0871	0.0744
7	0.0638	0.0767	0.0884	0.0949	0.0975	0.0982	0.0975	0.0949	0.0884	0.0767	0.0638
8	0.0558	0.0688	0.0795	0.0844	0.0860	0.0863	0.0860	0.0844	0.0795	0.0688	0.0558
9	0.0496	0.0626	0.0722	0.0758	0.0767	0.0769	0.0767	0.0758	0.0722	0.0626	0.0496
10	0.0446	0.0576	0.0660	0.0686	0.0692	0.0693	0.0692	0.0686	0.0660	0.0576	0.0446
11	0.0406	0.0534	0.0608	0.0626	0.0630	0.0630	0.0630	0.0626	0.0608	0.0534	0.0406
12	0.0372	0.0499	0.0563	0.0576	0.0577	0.0578	0.0577	0.0576	0.0563	0.0499	0.0372
13	0.0343	0.0469	0.0523	0.0532	0.0533	0.0533	0.0533	0.0532	0.0523	0.0469	0.0343
14	0.0319	0.0442	0.0488	0.0494	0.0495	0.0495	0.0495	0.0494	0.0488	0.0442	0.0319
15	0.0298	0.0418	0.0457	0.0462	0.0462	0.0462	0.0462	0.0462	0.0457	0.0418	0.0298
16	0.0279	0.0397	0.0430	0.0433	0.0433	0.0433	0.0433	0.0433	0.0430	0.0397	0.0279
17	0.0263	0.0378	0.0406	0.0408	0.0408	0.0408	0.0408	0.0408	0.0406	0.0378	0.0263
18	0.0248	0.0361	0.0384	0.0385	0.0385	0.0385	0.0385	0.0385	0.0384	0.0361	0.0248
19	.0235	0.0345	0.0364	0.0365	0.0365	0.0365	0.0365	0.0365	0.0364	0.0345	0.0235
20	0.0223	0.0330	0.0346	0.0347	0.0347	0.0347	0.0347	0.0347	0.0346	0.0330	0.0223

Table 3 $f(x)$ for $\delta=-0.5$ and $\delta=-0.4$

α \ x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	0.551	0.639	0.734	0.833	0.934	1.033	1.123	1.202	1.262	1.300	1.313
2	0.147	0.215	0.312	0.443	0.633	0.871	1.160	1.476	1.775	1.994	2.075
3	0.030	0.054	0.098	0.176	0.312	0.545	0.920	1.468	2.145	2.759	3.015
4	0.005	0.012	0.027	0.059	0.130	0.283	0.603	1.221	2.238	3.425	4.003
5	0.000	0.002	0.007	0.018	0.049	0.133	0.353	0.904	2.100	3.933	5.000
6	0.000	0.000	0.002	0.005	0.018	0.059	0.194	0.621	1.830	4.270	6.000
7	0.000	0.000	0.000	0.002	0.006	0.025	0.103	0.408	1.513	4.443	7.000
8	0.000	0.000	0.000	0.000	0.002	0.011	0.053	0.259	1.204	4.472	8.000
9	0.000	0.000	0.000	0.000	0.000	0.004	0.027	0.161	0.932	4.382	9.000
10	0.000	0.000	0.000	0.000	0.000	0.002	0.013	0.099	0.707	4.200	10.000

$\delta = -0.5$

α \ x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	0.597	0.685	0.778	0.872	0.964	1.049	1.122	1.178	1.213	1.226	1.213
2	0.181	0.263	0.378	0.533	0.736	0.977	1.244	1.496	1.680	1.748	1.680
3	0.042	0.076	0.136	0.242	0.423	0.714	1.139	1.665	2.141	2.340	2.141
4	0.009	0.019	0.043	0.094	0.205	0.437	0.885	1.622	2.483	2.902	2.483
5	0.002	0.005	0.012	0.034	0.091	0.242	0.618	1.436	2.690	3.420	2.690
6	0.000	0.001	0.004	0.012	0.039	0.126	0.404	1.191	2.778	3.904	2.778
7	0.000	0.000	0.000	0.004	0.010	0.064	0.254	0.943	2.769	4.363	2.769
8	0.000	0.000	0.000	0.001	0.006	0.032	0.156	0.724	2.688	4.808	2.688
9	0.000	0.000	0.000	0.000	0.003	0.016	0.094	0.543	2.553	5.244	2.553
10	0.000	0.000	0.000	0.000	0.001	0.008	0.056	0.401	2.384	5.677	2.384

$\delta = -0.4$

Table 4 $f(x)$ for $\delta = -0.3$ and $\delta = -0.2$

$\delta = -0.3$

α \ x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	0.649	0.737	0.826	0.913	0.993	1.062	1.116	1.149	1.161	1.149	1.116
2	0.231	0.332	0.469	0.645	0.859	1.093	1.315	1.477	1.537	1.477	1.315
3	0.064	0.115	0.204	0.356	0.602	0.961	1.404	1.865	1.973	1.865	1.404
4	0.016	0.035	0.078	0.170	0.363	0.735	1.347	2.061	2.409	2.061	1.347
5	0.004	0.010	0.028	0.076	0.201	0.513	1.152	2.233	2.855	2.233	1.152
6	0.000	0.003	0.010	0.032	0.166	0.339	0.958	2.379	3.272	2.379	0.958
7	0.000	0.000	0.003	0.014	0.055	0.216	0.803	2.357	3.713	2.357	0.803
8	0.000	0.000	0.001	0.006	0.028	0.135	0.627	2.327	4.165	2.327	0.627
9	0.000	0.000	0.000	0.002	0.014	0.083	0.479	2.291	4.623	2.291	0.479
10	0.000	0.000	0.000	0.000	0.007	0.050	0.360	2.138	5.092	2.138	0.360

$\delta = -0.2$

α \ x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	0.709	0.794	0.878	0.955	1.022	1.073	1.105	1.116	1.105	1.073	1.022
2	0.304	0.429	0.591	0.786	1.001	1.203	1.351	1.406	1.351	1.203	1.001
3	0.104	0.184	0.321	0.542	0.866	1.266	1.628	1.779	1.628	1.266	0.866
4	0.032	0.071	0.155	0.330	0.668	1.224	1.874	2.190	1.874	1.224	0.668
5	0.010	0.026	0.070	0.186	0.475	1.103	2.066	2.627	2.066	1.103	0.475
6	0.003	0.009	0.030	0.100	0.319	0.940	2.194	3.083	2.194	0.940	0.319
7	0.000	0.003	0.013	0.052	0.207	0.768	2.255	3.553	2.255	0.768	0.207
8	0.000	0.001	0.005	0.027	0.131	0.607	2.255	4.033	2.255	0.607	0.131
9	0.000	0.000	0.002	0.013	0.081	0.468	2.201	4.520	2.201	0.468	0.081
10	0.000	0.000	0.000	0.007	0.049	0.354	2.105	5.012	2.105	0.354	0.049

Table 5 $f(x)$ for $\delta = -0.1$ and $\delta = 0.0$

$\delta = -0.1$

α \ x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	0.776	0.858	0.933	0.998	1.048	1.080	1.091	1.080	1.048	0.998	0.933
2	0.407	0.561	0.747	0.950	1.143	1.283	1.335	1.283	1.143	0.950	0.747
3	0.174	0.304	0.514	0.820	1.199	1.542	1.685	1.542	1.199	0.820	0.514
4	0.068	0.148	0.316	0.640	1.174	1.796	2.099	1.796	1.174	0.640	0.316
5	0.025	0.068	0.180	0.461	1.072	2.007	2.552	2.007	1.072	0.461	0.180
6	0.009	0.030	0.098	0.313	0.923	2.154	3.027	2.154	0.923	0.313	0.098
7	0.003	0.013	0.052	0.205	0.760	2.230	3.514	2.230	0.760	0.205	0.052
8	0.001	0.005	0.027	0.130	0.603	2.240	4.007	2.240	0.603	0.130	0.027
9	0.000	0.002	0.013	0.081	0.466	2.193	4.505	2.193	0.466	0.081	0.013
10	0.000	0.000	0.007	0.049	0.353	2.101	5.002	2.101	0.353	0.049	0.007

$\delta = 0.0$

α \ x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	0.851	0.926	0.990	1.040	1.071	1.082	1.071	1.040	0.990	0.926	0.851
2	0.551	0.734	0.934	1.123	1.262	1.313	1.262	1.123	0.934	0.734	0.551
3	0.299	0.505	0.807	1.179	1.517	1.657	1.517	1.179	0.807	0.505	0.299
4	0.147	0.312	0.633	1.160	1.775	2.075	1.775	1.160	0.633	0.312	0.147
5	0.067	0.179	0.458	1.064	1.993	2.534	1.993	1.064	0.458	0.179	0.067
6	0.030	0.098	0.312	0.920	2.145	3.015	2.145	0.920	0.312	0.098	0.030
7	0.013	0.051	0.204	0.758	2.226	3.506	2.226	0.758	0.204	0.051	0.013
8	0.005	0.027	0.130	0.603	2.238	4.003	2.238	0.603	0.130	0.027	0.005
9	0.002	0.013	0.031	0.466	2.192	4.501	2.192	0.466	0.081	0.013	0.002
10	0.000	0.007	0.049	0.353	2.100	5.000	2.100	0.353	0.049	0.007	0.000

where μ is the mean. By removing the absolute sign from Equation (8), γ can be written as the sum of the two integrals,

$$\gamma(\alpha, \delta) = \int_{-\infty}^{\mu} (\mu - x) f(x) dx + \int_{\mu}^{\infty} (x - \mu) f(x) dx. \tag{9}$$

Nothing that μ is independent of the variable x and using relationships such as

$$\int_{-\infty}^{\mu} f(x) dx = F(\mu)$$

$$\int_{\mu}^{\infty} f(x) dx = 1 - F(\mu)$$

and

$$\int_{\mu}^{\infty} x f(x)dx = \mu - \int_{-\infty}^{\mu} x f(x)dx.$$

Equation (9) becomes

$$\gamma(\alpha,\ \delta) = 2\mu F(\mu) - 2\int_{-\infty}^{\mu} x f(x)dx. \tag{10}$$

Denote the integral in Equation (10) by $H(\mu)$ and for $\alpha > 0$,

$$\begin{aligned} H(\mu) &= \int_{-\infty}^{\mu} x f(x)dx \\ &= \int_{0}^{\mu} xG(\alpha,\ \delta) \frac{1}{\cosh^2[\alpha(x+\delta-0.5)]} dx \\ &= G(\alpha,\ \delta)\int_{0}^{\mu} \frac{x}{\cosh^2[\alpha(x+\delta-0.5)]} dx. \end{aligned}$$

The integral is similar to the one in the derivation of the mean with one exception, the upper limit is now μ instead to 1.0.

Evaluating this integral with proper limits, $H(\mu)$ becomes

$$H(\mu) = \frac{G(\alpha,\ \delta)}{\alpha^2} \alpha\mu \tanh[\alpha(\delta-0.5+\mu)] + \log \frac{\cosh[\alpha(\delta-0.5)]}{\cosh[\alpha(\delta-0.5+\mu)]}$$

for $\alpha > 0$ and becomes

$$H(\mu) = \int_{0}^{0.5} x dx = 0.125 = 0.125 \qquad \text{for } \alpha = 0.$$

Therefore $H(\mu)$ is

$$H(\mu) \begin{cases} = \dfrac{G(\alpha,\ \delta)}{\alpha^2} \alpha\mu \tanh[\alpha(\delta-0.5+\mu)] \\ \quad + \log \dfrac{\cosh[\alpha(\delta-0.5)]}{\cosh[\alpha(\delta-0.5+\mu)]} \\ \quad \text{for } \alpha > 0 \\ = 0.125 \quad \text{for } \alpha = 0. \end{cases} \tag{11}$$

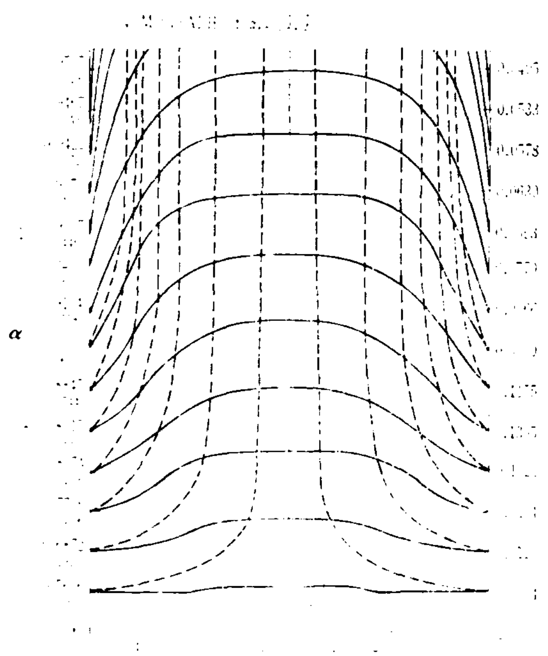


Fig. 1 Equi-mean curves (dotted lines), and equi-absolute deviation curves (solid lines). The numbers in parenthesis are the values of normalized absolute deviation, γ

The absolute deviation, γ, for the FRPDF is therefore

$$\gamma(\alpha,\ \delta) = 2\mu F(\mu) - 2H(\mu) \tag{12}$$

where the explicit expressions for μ, $F(\mu)$, and $H(\mu)$ are given in Equations (7), (4) and (11) respectively.

By use of Equations (7), (12) and the aid of the computer, the mean and the absolute deviation are computed for $\alpha = 0, 1, \cdots, 20$ and $\delta = -0.5, -0.4, \cdots, +0.4, +0.5$ and the equi-mean and equi-absolute deviation curves for $\alpha = 0, 1, \cdots, 10$ are given in the nomograph in Fig. 1.

By computing the mean and the absolute deviation of sample data, one can find the values of α and δ directly from the nomograph. Upon substituting these values of α and δ into Equation (4) one will obtain "the best fitted" probability distribution function for this particular

set of samples according to the criterion cited earlier (*i.e.*, this set of samples has the same mean and the same absolute deviation as that of the sample having the probability distribution function with the values of α and δ taken from the nomograph).

6. The Variance and Moment Generating Function

From a practical viewpoint the variance and moment generating function are of little use; however, they are investigated for theoretical considerations to verify that the variance is finite and that the moment generating function exists.

By definition the variance, σ^2, is:

$$\sigma^2=\int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx. \tag{13}$$

Equation (13) can be simplified as follows:

$$\sigma^2=\int_{-\infty}^{\infty}x^2 f(x)dx-\mu^2. \tag{14}$$

When $\alpha=0$, $f(x)=1$, $0\leq x\leq 1$ and $\mu=0.5$, therefore,

$$\sigma^2=\frac{1}{12}.$$

This is correct for the uniform distribution.

For $\alpha>0$, the integral in Equation (14) is difficult to evaluate and since it is shown that γ is used in lieu of σ^2, with the nomograph, for finding parameters α and δ, only the existence of the variance will be shown. Then from Equation (14)

$$\int_{-\infty}^{\infty}x^2 f(x)dx-\mu^2=G(\alpha,\ \delta)\int_0^1\frac{x^2}{\cosh^2[\alpha(x+\delta-0.5)]}dx-\mu^2.$$

Let $z=\alpha(x+\delta-0.5)$. Then

$$\int_{-\infty}^{\infty}x^2 f(x)dx=G\Big[\frac{1}{\alpha^3}\int_{\alpha(\delta-0.5)}^{\alpha(\delta+0.5)}\frac{z^2}{\cosh^2 z}dz+\frac{2(0.5-\delta)}{\alpha^2}\int_{\alpha(\delta-0.5)}^{\alpha(\delta+0.5)}\frac{z}{\cosh^2 z}dz$$
$$+\frac{(0.5-\delta)^2}{\alpha}\int_{\alpha(\delta-0.5)}^{\alpha(\delta+0.5)}\frac{1}{\cosh^2 z}dz\Big].$$

The first integrant can easily be shown to be integrable by use of the theorem: A function continuous on a closed interval is integrable there[6]. Also, since it is a proper distribution function, it has a unique characteristic function, by Levy's Theorem.

Let us denote the moment generating function for x by $M_x(\theta)$. By definition,

$$M_x(\theta)=\int_{-\infty}^{\infty}f(x)e^{\theta x}dx \tag{15}$$

When $\alpha=0$, $f(x)=1$, $0\leq x\leq 1$, therefore

$$M_x(\theta)=\int_0^1 e^{\theta x}dx=\frac{e^\theta-1}{\theta}.$$

When $\alpha>0$, Equation (15) becomes,

$$M_x(\theta)=\int_0^1 G(\alpha,\ \delta)\frac{1}{\cosh^2[\alpha(x+\delta-0.5)]}\cdot e^{\theta x}dx$$
$$=G(\alpha,\ \delta)\int_0^1\frac{e^{\theta x}}{\cosh^2[\alpha(x+\delta-0.5)]}dx.$$

Using the same argument of continuity and integrability of this integrant, it follows that the moment generating function $M_x(\theta)$ exists.

To facilitate application of the FRPDF, the following tables and a Nomogram is provided. Also, several figures showing the shape of the FRPDF with different values of parameters α and δ are given. The attached Appendix will help the reader get a first-hand feel

of its ease in application and goodness-of-fit.

7. *Conclusions*

The objective of this paper has been to present a finite range probability distribution function which was developed by Braswell and further refined and applied by Manders. The probability density function, $f(x)$, is strictly unimodal as shown in Figs. 2 through 7. When $\delta = 0$, the probability density function is always symmetrical with respect to the vertical line $x = 0.5$, and the distribution function is a family of S-shaped curves. When $\delta < 0$, the peak of the probability density function shifts toward the $x = 1$ line and when $\delta > 0$, it shifts to the $x = 0$ line. When $\alpha = 0$, the FRPDF distribution is identically uniform; for the large value of α, $\alpha = 10$ or more, the peak of the distribution becomes narrower and higher at appropriate value of x for given value of δ. Eventually, this peak will reach infinity in the limit, and the function becomes an impulse function at a certain value of x. This is very convenient in practical cases. Since this indicates that for a certain statistical sample, if the value of α is large, this statistical sample can be treated approximately as a deterministic one. Hence δ can be considered as the location parameter, while α can be considered as the shape parameter.

The new FRPDF can be easily applied to almost any practical problems where experimental data are easily collected. Once the necessary data are tabulated, one can easily compute the mean and the absolute deviation of that sample data. Then by use of the Nomogram in Fig. 1 the values of the parameters, α and δ, can be found directly. Substituting these

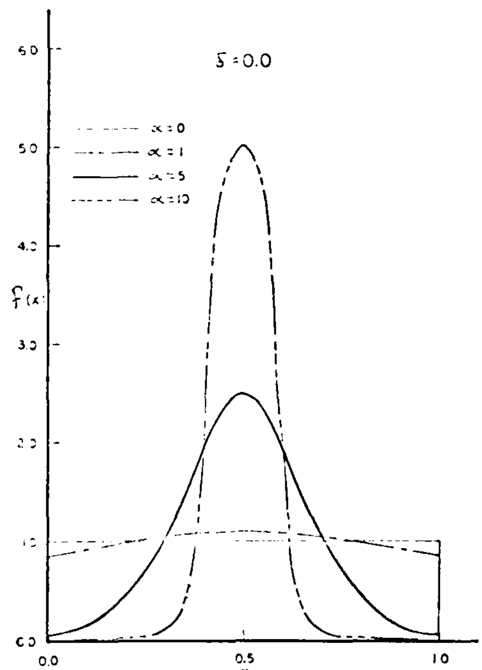


Fig. 2 $f(x)$ for the FRPDF $\delta = 0.0$

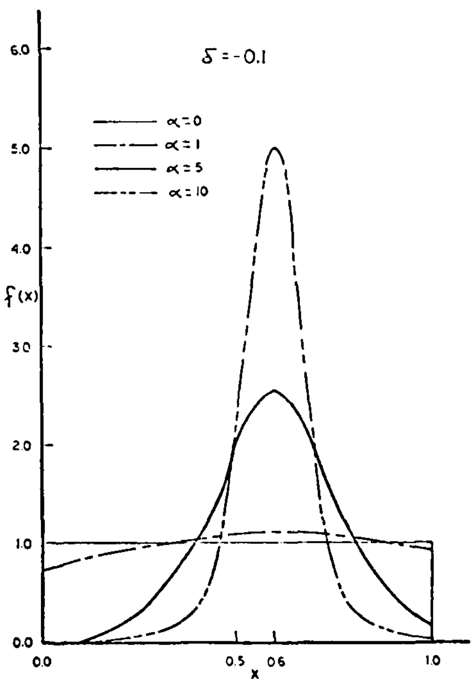


Fig. 3 $f(x)$ for the FRPDF $\delta = -0.1$
If $\delta = +0.1$, replace x by $1 - x$

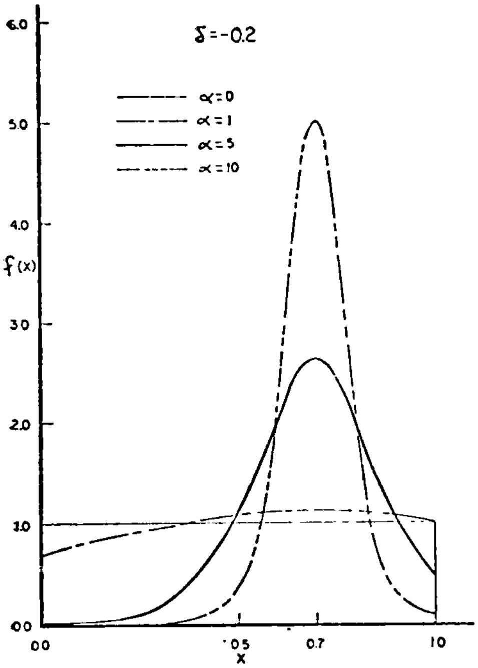


Fig. 4 $f(x)$ **for the FRPDF** $\delta=-0.2$
If $\delta=+0.2$, **replace** x **by** $1-x$

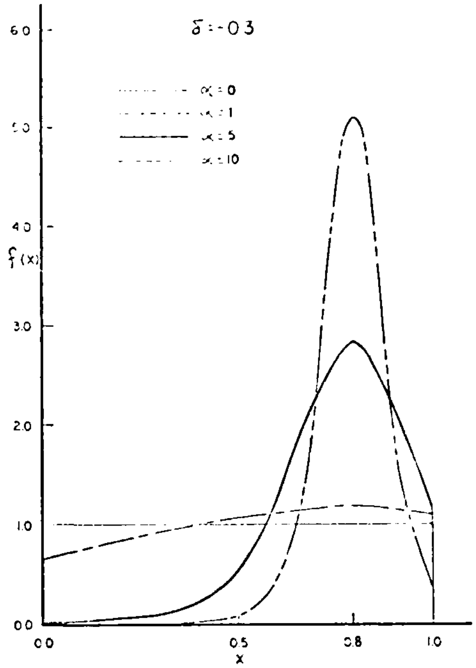


Fig. 5 $f(x)$ **for the FRPDF** $\delta=-0.3$
If $\delta=+0.3$, **replace** x **by** $1-x$

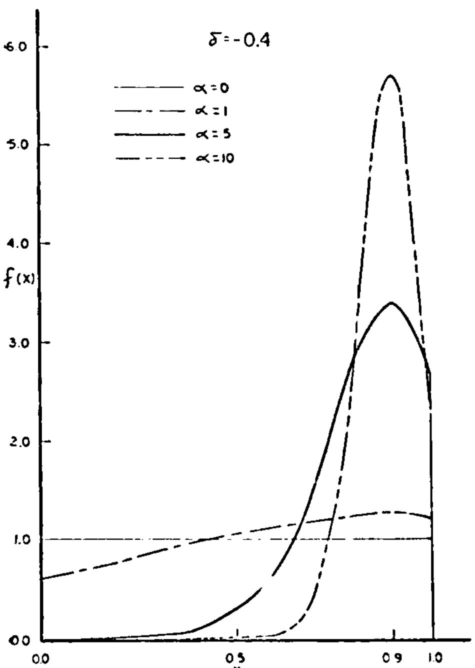


Fig. 6 $f(x)$ **for the FRPDF** $\delta=-0.4$
If $\delta=+0.4$, **replace** x **by** $1-x$

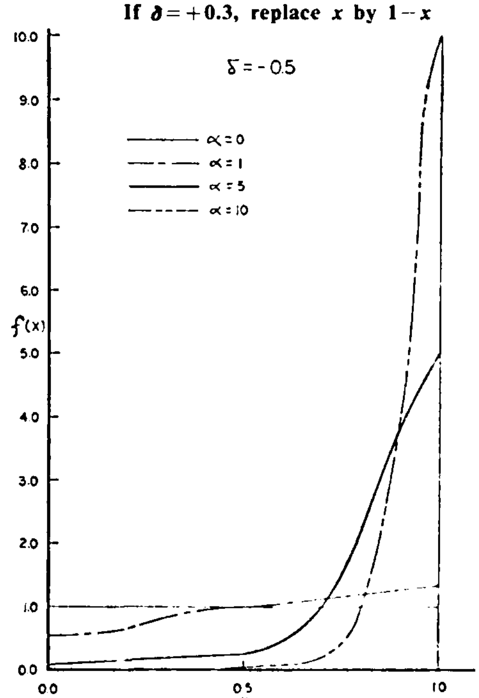


Fig. 7 $f(x)$ **for the FRPDF** $\delta=-0.5$
If $\delta=+0.5$, **replace** x **by** $1-x$

values of α and δ into Equation (4) one can obtain "the best fitted" probability distribution function.

REFERENCES

[1] Pearl, R. (1930): *The Biology of Population Growth*, Alfred A. Knopf, Inc., New York.

[2] Braswell, R. N. (1969): "On Modeling the Resource Allocation and Control Problem," *Proceedings of the American Astronautical Society*, June, pp. 11.

[3] Papoulis, A. (1965): *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill Book Co., New York, pp. 92.

[4] Dwight, H. B. (1957): *Tables of Integrals and Other Mathematical Data*, MIT Press, 3rd Ed., pp. 150.

[5] Special credit is due Dr. Arnfinn Manders, University of Florida, Genesys Center, Port Canaveral, Florida for his helpful suggestions on using the absolute deviation to approximate a measure of dispersion.

Also, special credit is due Dr. A. K. Varma, University of Florida for his help on mathematical theory and notation.

[6] Sokolnikoff, I. S. (1939): *Advanced Calculus*, McGraw-Hill Book Co., Inc., New York.

Note: The use of the New Probability Distribution Function is illustrated in Vol. 17, No. 2, 1970, section B of this journal, entitled "On testing and Application of A New Finite Range Probability Distribution Function."

ON TESTING AND APPLICATION OF A NEW FINITE RANGE PROBABILITY DISTRIBUTION FUNCTION*

By Robert N. BRASWELL
University of Florida
Gainesville, Florida

Clara Fu-Mei MANDERS
Radiation, Incorporated Melbourne, Florida

To demonstrate flexibility and sensitivity this paper will cover some aspects of testing and application. There have been several other experiments using the new FRPDF and in all cases it has proven superior in results and much easier to use. Even when the distribution of the data was known the FRPDF was preferable. The complex expression did not deter new users from being attracted to further applications.

1. Testing and Application

An important problem in statistics is to find how well a sample taken from a population agrees with some distribution function assumed for that population. Two such tests are considered here.

The range of x is divided into M equal regions and the number of sample points falling within each region is counted. Let $Y_1, Y_2, \cdots, Y_M$ be the result. From the assumed distribution and the size of the sample, the expected number of points in each region is computed: $g_1, g_2, \cdots, g_M$. The deviation between this and the actual result is expressed by

$$D=\sum_{m=1}^{M}\frac{(y_m-g_m)^2}{g_m} \tag{1}$$

where

$g_m = Np_m$
= the number of expected points in the m-th interval

$N=\sum_{m=1}^{M} y_m$
= sample size

y_m = the number of sample points in the m-th interval

$$p_m=\int_{x_{m-1}}^{x_m} f(x)dx = F(x_m)-F(x_{m-1}) \qquad m=1, \cdots, M$$

$$F(x_{m-1})=0 \qquad \text{for } m=1$$

the probability of sample points falling in the m-th interval.

This deviation is used to ascertain the confidence level of the assumed distribution.

* Received 24. Feb., 1970.
(Reprinted from Rep. Stat. Appl. Res., JUSE, Vol 17, No. 2, 1970, by permission of the publisher.)

As an application of the method described in this paper, consider a class grade distribution of 50 graduate engineering students. Table 1 gives the grades for the students [1].

STEP I Compute an estimate of the normalized mean, $\hat{\mu}$, and the absolute deviation, $\hat{\gamma}$, from the data given in Table 1. For convenience in analysis we will normalize the grade range from [0, 100] to [0, 1]. They are found to be;

$$\hat{\mu}=0.829$$
$$\hat{\gamma}=0.093.$$

Table 1 Class grades of a graduate engineering class of 50 students

i	*Gr.*	*i*	*Gr.*	*i*	*Gr.*	*i*	*Gr.*	*i*	*Gr.*
1	81	11	85	21	80	31	67	41	78
2	86	12	69	22	49	32	89	42	95
3	78	13	95	23	96	33	77	43	86
4	85	14	87	24	63	34	95	44	75
5	79	15	69	25	92	35	63	45	96
6	85	16	85	26	66	36	100	46	96
7	92	17	90	27	52	37	85	47	98
8	83	18	71	28	79	38	97	48	83
9	96	19	68	29	94	39	85	49	81
10	89	20	79	30	96	40	90	50	91

STEP II Locate the intersection of $\hat{\mu}=0.829$ curve with $\hat{\gamma}=0.093$ curve on Fig. 1, and estimate the coordinates of this intersection.

This is equivalent to the estimation of two parameters α and δ in our $F(x)$. Here

$$\hat{\alpha}=5.5$$
$$\hat{\delta}=-0.4.$$

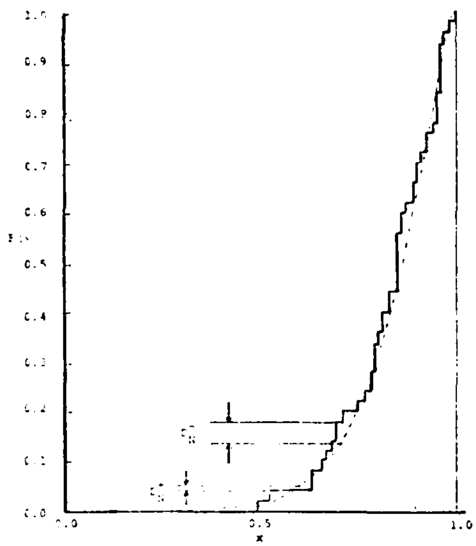


Fig. 1 The FRPDF $F(x)$ with $\alpha=5.5$, $\delta=-0.4$. Fitted to curve of class grade

STEP III Substituting these values into Equation (4), we obtain

$$F(x)=\frac{\cosh(0.55)\sinh(5.5x)}{\sinh(5.5)\cosh[5.5(x-0.9)]}. \quad (2)$$

STEP IV Divide the range of x into 10 equal regions: $M=10$

1) Compute $F(x_m)$:

$$F(x_m)=\frac{\cosh(0.55)\sinh(5.5x_m)}{\sinh(5.5)\cosh 5.5(x_m-0.9)}$$
$$m=1, \cdots, 10.$$

With the aid of the computer, the values of $F(x_m)$ are, 0.0001, 0.0005, 0.0017, 0.0054, 0.0161, 0.0473, 0.1329, 0.3328, 0.6664, and 1.000.

2) Compute p_m:

$$p_m=\int_{x_{m-1}}^{x_m} f(x)dx=F(x_m)-F(x_{m-1}),$$

$$m = 1,\ 2,\ \cdots,\ M.$$

Using the values for $F(x_m)$, the values of P are, 0.0001, 0.0004, 0.0012, 0.0037, 0.0107, 0.0312, 0.0856, 0.1999, 0.3336, and 0.3336.

STEP V Prepare Table 2.

Table 2 Distribution of class grades

m	y_m	$g_m = Np_m$	$\frac{(g_m - y_m)^2}{g_m}$	m	y_m	$g_m = Np_m$	$\frac{(g_m - y_m)^2}{g_m}$
1	0	0.005	0.005	7	7	4.280	1.729
2	0	0.020	0.020	8	9	9.995	0.099
3	0	0.060	0.060	9	17	16.680	0.006
4	0	0.185	0.185	10	15	16.680	0.169
5	1	0.535	0.404				
6	1	1.560	0.201	Total	50	50	2.878

2. The χ^2-Test [2, 3]

The purpose of the χ^2 test is to find the probability that the observed deviation between the theoretical model and the empirical data is in fact due to the random nature of this set of data.

If the y_m are sufficiently large, say more than 10, the deviation we defined in (1), is distributed according to the χ^2 distribution with $M-1$ degrees of freedom. This is K. Pearson's test function which gives great weight to those deviation squares $(y_m - g_m)^2$ that correspond to small p_m. If the assumed parent distribution is not completely known and k parameters defining it have been determined to fit the sample, the number of degrees of freedom is reduced to $M-1-k$.

Now we will apply the χ^2-test to our example.

$$\chi^2 = \sum_{m=1}^{M} \frac{(g_m - y_m)^2}{g_m} \qquad (3)$$

$$= 2.878$$

and the degrees of freedom are 7 $(10-1-2)$.

For 7 degrees of freedom, this deviation is exceeded about 90 percent of the time [3]. The assumption of our FRPDF is therefore very good. There is thus nothing in the value of χ^2 to lead us to reject our hypothesis.

3. The Kolmogorov-Smirnov Test [3]

It is also desired to investigate how well our empirical data fits our theoretical distribution by $k-s$ test. The $k-s$ test allows us to place confidence level on the positive as well as negative deviations, *i.e.*, it allows us to check the theoretical distribution for points of excessive as well as inadequate probability.

Define the one-sided deviations as:

$$D_N^+ = \sup_{0 \le x \le 1} [F(x_m) - S_m(x_m)] \quad \text{and} \quad D_N^- = \sup_{0 \le x \le 1} [S_m(x_m) - F(x_m)] \qquad (4)$$

where

$$S_m(x_m) = \frac{\sum_{i=1}^{m} y_i}{N} \quad \text{and} \quad F(x_m)$$

are given in Step IV.

According to Smirnov's asymptotic distribution [3],

$$P\{D_N^+ > \varepsilon\} \sim e^{-2N\varepsilon^2} \quad \text{and} \quad P\{D_N^- > \lambda\} \sim e^{-2N\lambda^2}. \tag{5}$$

Now we proceed to test the example.

$$D_N^+{}_{\text{empirical}} = 0.0014 \quad \text{and} \quad D_N^-{}_{\text{empirical}} = 0.0471.$$

Table 3 Calculation of the one-sided deviations, D_N^+ and D_N^-

m	$F(x_m)$	$S_m(x_m)$	D_N^+	D_N^-	m	$F(x_m)$	$S_m(x_m)$	D_N^+	D_N^-
1	0.0001	0.00			6	0.0473	0.04	0.0073	
2	0.0005	0.00			7	0.1329	0.18		0.0471
3	0.0017	0.00			8	0.3328	0.36		
4	0.0054	0.00			9	0.6664	0.70		
5	0.0161	0.02			10	1.0000	1.00		

See Fig. 1 for comparisons.

4. The Glass Bottom Boat Problem

Consider the "glass bottom boat problem" which is similar to the well known "Newsboy Problem." [4]

Statement of the Problem:

A photographer must decide how many pictures to print each time after he takes a picture of the passengers in the glass bottom boat at Silver Springs, Florida. Suppose that the total number of the passengers is fixed, say forty. The cost of a picture is C and the selling price is S. Any pictures not sold at the end of the day are a total loss. Let $p(y)$ be the probability that y pictures will be demanded each time. Then his expected profit for each boat if he prints h pictures is

$$E[P(h)] = S\sum_{y=0}^{h} yp(y) + Sh\Big[\sum_{y=h+1}^{\infty} p(y)\Big] - Ch$$

since the revenue received is Sy if $y \leq h$, and is Sh if $y > h$. The problem is to determine the value of h which maximizes his expected profit.

Often it is convenient to treat h and the demand variable y as continuous. Then if $f(y)$ is the density function for demand and $F(y)$ is its distribution function, the expected profit for each boat when h units are printed is

$$E[P(h)] = S\int_0^h yf(y)dy + Sh\int_h^{\infty} f(y)dy - Ch. \tag{6}$$

The optimal h is then a solution to $dE[P(h)]/dh = 0$.

Using Leibnitz's Rule, we obtain

$$\frac{dE}{dh} = 0 = S - C - SF(h).$$

Thus the optimal h, h^*, satisfies the equation

$$F(h) = \frac{S-C}{S}. \tag{7}$$

Let $S = \$1.20$ and $C = \$0.50$. Then $F(h) = 7/12$, or 0.583.

Equation (6) is a strictly concave function of h. This implies that any relative maximum of $E[P]$ is the absolute maximum and the absolute maximum is unique.

Suppose the demands are normally distributed then Equation (7) becomes

$$\Phi\left(\frac{h-\bar{y}}{\sigma_y}\right) = \frac{S-C}{S} \tag{8}$$

where

$\bar{y}$ is the sample mean, and

σ_y is the standard deviation.

If the demands follow the FRPDF then Equation (B.7) becomes

$$F(h_r) = \frac{\cosh[a(\delta+0.5)]}{\sinh(a)} \frac{\sinh(ah_r)}{\cosh[a(h_r-\delta-0.5)]} = \frac{S-C}{S} \tag{9}$$

where h_r is normalized h.

In order to determine the demand the photographer performs the following experiment. For each of ten successive boat loads he prints forty pictures (the maximum possible demand). He then records the number of prints that he sells to each boat load of passengers. In this manner he obtains the following table:

Table 4 Number of pictures sold to each boatload of people

Boat Number i	Number of Demands y_i	Boat Number i	Number of Demands y_i
1	36	6	30
2	34	7	25
3	39	8	37
4	20	9	23
5	17	10	33

Table 5 Calculation of $\bar{y}$, $|y_i-\bar{y}|$ and $(y_i-\bar{y})^2$

i	y_i	$\|y_i-\bar{y}\|$	$(y_i-\bar{y})^2$	i	y_i	$\|y_i-\bar{y}\|$	$(y_i-\bar{y})^2$
1	36	6.6	43.56	7	25	4.4	19.36
2	34	4.6	21.16	8	37	7.6	57.76
3	39	9.6	92.16	9	23	6.4	40.96
4	20	9.4	88.36	10	33	3.6	12.96
5	17	12.4	153.76				
6	30	0.6	.36	Total	294	65.2	530.40

We find that the estimator for μ_y is the sample mean $\bar{y}$,

$$\bar{y} = \frac{1}{10}\sum_{i=1}^{10} y_i = 29.4.$$

The estimator for σ_y^2 is:

$$\sigma_y^2 = \frac{\sum_{i=1}^{10}(y_i-\bar{y})^2}{10-1} = 58.93$$

and that for σ_y is.

$$\sigma_y = \sqrt{58.93} = 7.68.$$

The estimator for γ, is the sample absolute deviation $\overline{|y-\bar{y}|}$,

$$\overline{|y-\bar{y}|} = \frac{1}{10}\sum_{i=1}^{10}|y_i-\bar{y}|=6.52.$$

Now we are in the position to calculate h^* by assuming that $F(h)$ in Equation (7) is 1) normal distribution function and 2) FRPDF.

1) Assume $F(h)$ is normal distribution function. Then h^* is the solution to Equation (8), or

$$\Phi\left(\frac{h-29.4}{7.68}\right)=0.583.$$

Hence from the normal tables

$$\frac{h-29.4}{7.68}=0.21 \quad \text{or} \quad h=31.$$

Thus 31 pictures should be printed.

2) Assume $F(h)$ is FRPDF. In this model the variable is normalized; we should make a liner transformation on each sample value.

We find the normalize $\bar{y}$, $\hat{\mu}$ to be:

$$\hat{\mu}=\frac{29.4}{40}=0.735$$

and the normalized $\overline{|y-\bar{y}|}$, $\hat{\gamma}$ to be

$$\hat{\gamma}=\frac{6.52}{40}=0.163.$$

From Fig. 1, we can estimate the parameters of our FRPDF α and δ. They are found to be

$$\hat{\alpha}=3$$
$$\hat{\delta}=-0.4.$$

Table 6 Calculation of the expected profit for three different decisions

i	y_i	Number of Pictures Sold		
		$h=y=29$	$h=31$ (Normal)	$h=33$ (FDPDF)
1	36	29	31	33
2	34	29	31	33
3	39	29	31	33
4	20	20	20	20
5	17	17	17	17
6	30	29	30	30
7	25	25	25	25
8	37	29	31	33
9	23	23	23	23
10	33	29	31	33
Total sold		259	270	280
Total printed		290	310	330
Total return		\$310.80	\$324.00	\$336.00
Total cost		145.00	155.00	165.00
Total profit		165.80	169.00	171.00

Substituting these values into Equation (9) and rewriting it, we have

$$\coth(3h_0) = -\tanh(-2.7) + \frac{12}{7}\,\frac{\cosh(0.3)}{\cosh(-2.7)\sinh(3)}$$

$$= 0.99101 + \frac{12}{7}\,\frac{(1.0453)}{(7.4735)(10.018)}$$

$$= 1.01494$$

or

$$\tanh(3h_0) = 0.98528$$

$$3h_0 = 2.46$$

$$h_0 = \frac{2.46}{3} \quad \text{and} \quad h = \frac{2.46}{3} \times 40 = 33.$$

Thus 33 pictures should be printed in this case.

Of the three methods employed we see that the new FRPDF gives the BEST decision as to number of pictures to be printed. More involved experiments with the Newsboy Problem, the Glass Bottom Boat Problem, etc., yield comparable results [6].

REFERENCES

[1] Special Credit is due Dr. R. S. Leavenworth, University of Florida, for permiting his ISE 621 class to be used for experimental purposes.

[2] Kendall, M. G. (1948): *The Advanced Theory of Statistics*, 4th Ed., **1**, Charles Griffin & Co., Ltd., pp. 290-301.

[3] Von Mises, R. (1964): *Mathematical Theory of Probability and Statistics*, Academic Press Inc., New York, pp. 447-449, pp. 490-492.

[4] Hald, A.: *Statistical Tables and Formulas*, John Wiley & Sons, Inc., New York, pp. 40.

[5] Hadley and Whitin (1963): *Analysis of Inventory Systems*, Prentice Hall, Inc., Englegood Cliffs, N.J., pp. 297-299.

[6] Special Credit is due Dr. Arnfinn Manders, University of Florida, Genesys Center, Port Canaveral, Florida for his helpful suggestions on the testing procedures.

(The reverse of this page is blank)